



HAL
open science

La polysémie régulière : approches empiriques. Présentation

Lucie Barque, Pauline Haas

► **To cite this version:**

Lucie Barque, Pauline Haas. La polysémie régulière : approches empiriques. Présentation. *Lexique*, 2022, 31, pp.5-25. halshs-03894572

HAL Id: halshs-03894572

<https://shs.hal.science/halshs-03894572>

Submitted on 12 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La polysémie régulière : approches empiriques

Présentation

Lucie Barque

LLF, UMR 7110 (CNRS / Université Paris Cité)

& Université Sorbonne Paris Nord

lucie.barque@univ-paris13.fr

Pauline Haas

Lattice, UMR 8094 (CNRS / ENS-PSL / Paris 3)

& Université Sorbonne Paris Nord

pauline.haas@univ-paris13.fr

1. Introduction*

La polysémie régulière, qui désigne les alternances de sens récurrentes au sein de classes lexicales (Apresjan, 1974), a fait l'objet d'un nombre important de travaux théoriques. Ces derniers se sont concentrés essentiellement sur deux grandes questions : (i) celle de savoir quel type de représentation linguistique adopter pour les mots présentant des alternances de sens régulières et (ii) celle de savoir comment décrire ces alternances de sens et leur récurrence dans le lexique (pour un aperçu général, voir entre autres Ravin & Leacock, 2000 ; Falkum & Vicente, 2015 ; Dölling, 2020)¹. Plus récemment, des études empiriques s'appuyant

* Nous remercions chaleureusement les auteurs, les relecteurs et les éditrices de la revue *Lexique* d'avoir participé à l'élaboration de ce numéro.

¹ Bâties sur la critique de l'approche discrète de la polysémie, qui consiste à établir la liste des différents sens possibles d'un mot, deux grandes alternatives ont été proposées. L'approche sous-spécifiée, d'inspiration cognitiviste, qui met en doute la possibilité de lister l'ensemble des sens et effets de sens possibles d'un mot, défend l'idée d'une représentation linguistique unifiée qui n'encode que l'information sémantique commune à l'ensemble des usages de l'unité polysémique (Victorri & Fuchs, 1996). L'approche générative, qui critique principalement la non prise en compte par l'approche discrète des régularités polysémiques, propose tantôt des descriptions uniques mais complexes et structurées permettant de cibler le sens d'une unité en contexte aux moyens de mécanismes compositionnels (Pustejovsky, 1995), tantôt des représentations linguistiques primaires à partir desquelles sont générées des sens au moyen de règles d'extension soumises à un certain nombre de contraintes d'application, lexicales ou pragmatiques (Lehrer, 1990 ; Ostler & Atkins, 1991 ; Nunberg & Zaenen, 1992 ; Copestake & Briscoe, 1995 ; entre autres).

sur des données linguistiques primaires (données comportementales, corpus textuels) et/ou secondaires (jugements, ressources lexicales) ont produit des résultats susceptibles d'apporter des éclairages nouveaux sur ces questions théoriques, en mettant notamment en lumière le caractère hétérogène et fondamentalement graduel du phénomène observé. L'objectif de cette introduction du numéro spécial sera de présenter, parmi ces études, celles qui cherchent à mesurer la variation des polysémies observées en termes de régularité ou qui s'intéressent aux facteurs explicatifs de cette variation.

1.1. Définitions préliminaires

Avant de nous intéresser à la polysémie régulière proprement dite, rappelons la définition communément admise de la polysémie et des deux notions qui lui sont complémentaires, l'homonymie d'une part, la monosémie d'autre part. On dit d'un mot qu'il est polysémique s'il a plusieurs sens distincts mais liés entre eux (Bréal, 1904 ; Kleiber, 1999 ; entre autres). C'est le cas par exemple du nom *bureau*, qui peut dénoter un 'meuble' ou une 'pièce', ces deux sens étant unis par un lien clair : le meuble et la pièce sont dédiés à la même activité, à savoir le travail. La condition de sens distincts trace une première frontière entre polysémie et monosémie, qui caractérise les mots dotés d'un sens unique². La condition de l'existence d'un lien entre les différents sens du mot polysémique dessine une seconde frontière avec l'homonymie, qui implique des sens distincts mais sans relation entre eux. Si d'un point de vue définitionnel les frontières entre homonymie, polysémie et monosémie sont nettes, l'évaluation de ces conditions peut s'avérer difficile, lorsque le consensus est faible sur l'existence d'un lien entre les sens d'un mot (ex. *jalousie* 'sentiment' / 'persienne'), ou lorsque le caractère distinct des sens n'apparaît pas clairement (ex. *amour* 'sentiment amoureux' / 'amitié'). Les frontières entre les catégories étant poreuses, il a été proposé de voir la polysémie comme occupant une place centrale le long d'un continuum entre homonymie et monosémie (Tuggy, 1993 ; Victorri & Fuchs, 1996 ; entre autres).

Une polysémie est dite régulière si l'on observe au moins deux mots dans la langue qui présentent une même extension de sens (Apresjan, 1974), elle est irrégulière dans le cas

² Certains mots monosémiques peuvent présenter des variations de sens mais sans que ces variations soient perçues comme mettant en jeu des sens distincts. Un exemple fréquemment donné est celui des noms de parenté, telle que *tante* par exemple, qui peut désigner aussi bien la sœur du père que celle de la mère (Tuggy, 1993) mais sans que ces deux sens (lexicalisés par des mots de formes différentes dans certaines langues), soient perçus comme distincts par les locuteurs du français, la différence étant purement référentielle.

contraire. Par exemple, la polysémie du nom *bureau* évoquée plus haut est régulière dans le lexique du français puisqu'on trouve au moins un autre nom présentant la même alternance entre un meuble et un lieu (ex. *bibliothèque, bar*). À l'inverse, la polysémie du nom *mètre*, qui unit les sens 'unité de mesure' (ex. *deux mètres cinquante*) et 'instrument de mesure' (ex. *prête-moi ton mètre*) est irrégulière, puisqu'il n'existe a priori pas d'autre nom en français contemporain qui présente la même alternance de sens. La frontière entre polysémie régulière et polysémie irrégulière dépend toutefois crucialement de deux paramètres : le grain sémantique choisi pour décrire l'alternance sémantique, d'une part, et l'étendue du lexique considéré, d'autre part. Le lien métonymique irrégulier entre les deux sens du nom *mètre* pourrait ainsi être considéré comme régulier si l'on généralise le sens source à celui de 'quantité' – le nom *mesure* relèverait alors du lien décrit (ex. *deux mesures d'eau / lots de mesures en étain*) – ou si l'on étend l'évaluation au lexique spécialisé ou diachroniquement marqué – le lien s'appliquerait alors également aux noms *toise* (ex. *deux toises plus bas / une toise en bois*) et *litre* (ex. *deux litres de lait / un litre en bois*).

L'existence de polysémies de faible régularité n'a que peu attiré l'attention des auteurs (Pethö, 2007 ; Brocher *et al.*, 2018), qui se sont intéressés principalement aux polysémies clairement régulières et aux polysémies systématiques. Les premières concernent une proportion notable des effectifs d'une classe lexicale, sans toutefois les concerner tous. C'est le cas par exemple de la polysémie "action > objet résultant de cette action" qui est régulière (ex. *traduction, construction, sculpture*) mais qui ne concerne qu'une partie des noms d'action (par ex. le nom *réparation* n'a pas de sens d'objet résultant). Les polysémies systématiques, quant à elles, s'appliquent à l'ensemble d'une classe lexicale donnée, fut-elle très restreinte (Apresjan, 1974). Les noms de contenants (ex. *seau, fourchette, sachet*), par exemple, peuvent systématiquement s'employer pour désigner une quantité mesurée par ce contenant (*deux seaux d'eau, trois fourchettes de riz, dix sachets de cocaïne*). Là encore, le grain adopté pour décrire les conditions sémantiques d'une extension de sens peut jouer sur l'évaluation de sa régularité : plus le grain est fin, plus la proportion des noms polysémiques au sein d'une classe est importante, jusqu'à arriver théoriquement à une systématité, ce qui soulève la question de l'existence ou non d'une différence de nature entre une polysémie régulière non systématique et une polysémie régulière systématique.

Terminons ces remarques préliminaires par une précision concernant le périmètre de notre objet d'étude. Le phénomène de la polysémie systématique, au sens où elle est définie ci-dessus, est souvent confondu avec celui de la monosémie complexe. Ce phénomène, décrit dans la littérature sous les termes de *facettes* (Cruse, 1995), de *types pointés* (Pustejovsky,

1995), voire de *polysémie inhérente* ou *logique* (Copestake & Briscoe, 1995 ; Frisson, 2015), désigne des unités qui dénotent des référents complexes, et dont les différents aspects les font relever de classes sémantiques distinctes mais liées entre elles, les faisant ressembler en cela à des cas de polysémie. L'exemple classique est celui des noms de documents (*livre, catalogue, magazine, etc.*), qui désignent des contenus informationnels sur supports et relèvent donc de la classe sémantique des objets informationnels et de celle des objets physiques. Le trait principal qui distingue polysémie systématique et monosémie complexe a été largement décrit dans la littérature : les sens concernés par une polysémie régulière sont mutuellement exclusifs en contexte, au contraire des facettes des monosèmes complexes, ce que révèle en particulier le test de la copredication (Zwicky & Sadock, 1975 ; Cruse, 1986 ; Godard & Jayez, 1996 ; Asher, 2011 ; Moldovan, 2021 ; en autres). Comme d'autres avant nous (par ex. Dölling, 2020), nous prendrons soin ici de bien distinguer ces deux phénomènes pour nous concentrer sur la seule polysémie régulière³ dont le spectre s'étend donc de la polysémie minimalement régulière à la polysémie systématiquement régulière.

1.2. La polysémie régulière, un phénomène graduel

Pourquoi chercher à évaluer précisément le degré de régularité d'une polysémie ? On peut penser tout d'abord que la représentativité effective d'une polysémie dans la langue a une incidence sur l'acquisition des sens produits par cette polysémie. Des études montrent un avantage de la polysémie par rapport à l'homonymie sur cette question : les enfants acquièrent plus facilement un sens nouveau d'un mot existant quand il est lié à l'un des sens établis de ce mot que quand il ne l'est pas (Srinivasan *et al.*, 2019) et le stockage d'un sens nouvellement acquis entre moins en compétition avec un autre sens existant lorsque ceux-ci sont sémantiquement liés (Rodd *et al.*, 2012 ; Maciejewski *et al.*, 2020). On peut donc faire l'hypothèse que plus le sens nouvellement créé est lié de manière régulière au sens existant (et donc plus souvent on a été confronté à ce type de polysémie), plus l'acquisition en est facilitée et moins est forte la compétition⁴ entre le sens nouvellement intégré au lexique

³ Nous signalerons toutefois, le cas échéant, les études qui incluent dans leurs données des cas de monosémie complexe.

⁴ La compétition s'observe dans le cadre d'une tâche de décision lexicale qui consiste, pour les participants, à dire si deux mots qui leur sont proposés sont liés sémantiquement (réponse par « oui » ou « non » après que la paire apparaît sur l'écran). Le second mot de la paire correspond au mot cible, un mot initialement monosémique pour lequel un nouveau sens a été créé pour les besoins de l'expérience. Certains de ces sens nouveaux sont liés au sens existant du mot par une extension, d'autres

mental et le sens stocké de plus longue date. La caractérisation précise du degré de régularité des types de polysémie pourrait ainsi faire avancer notre connaissance de la structuration du lexique mental et de la représentation des mots à sens multiples. Elle pourrait également permettre, d'un point de vue typologique, une comparaison plus fine entre les langues. Certaines polysémies régulières s'observent dans différentes langues, avec des effectifs (i.e. l'ensemble des mots présentant cette polysémie) plus ou moins comparables, comme l'ont montré de nombreuses études de cas, par exemple celles portant sur la métaphore "animal > personne" (Talebinejad & Dastjerdi, 2005 ; Rodriguez, 2009 ; entre autres). Des études à plus large spectre, menées à partir de ressources multilingues (Peters, 2003) ou de questionnaires soumis à des locuteurs de différentes langues (Srinivasan & Rabagliati, 2015) laissent par ailleurs entrevoir l'existence d'un lien entre régularité d'une polysémie et récurrence de cette polysémie à travers les langues (voir *infra*, section 3).

L'hypothèse d'une corrélation entre régularité d'une polysémie et facilité d'acquisition d'une part et tendance à l'universalité d'autre part implique, pour être vérifiée, qu'on soit à même d'évaluer précisément le degré de régularité d'une polysémie donnée, ce qui soulève un certain nombre de questions aussi bien théoriques que pratiques (Barque *et al.*, 2018). Dans la suite de la présentation, nous nous intéresserons tout d'abord aux études qui ont cherché à évaluer le degré de régularité d'une polysémie en calculant ses effectifs dans la langue (section 2). Nous nous pencherons ensuite sur les sources de la polysémie régulière en évoquant des travaux qui décrivent de potentiels facteurs explicatifs des degrés de régularité des polysémies observées, parmi lesquels la figure sur laquelle repose l'extension de sens (métonymie / métaphore), le degré de connectivité entre les sens en jeu, et enfin la construction morphologique des mots ayant ces sens (section 3). Nous terminerons en présentant les quatre contributions de ce numéro spécial consacré aux approches empiriques de la polysémie régulière (section 4).

n'ont aucun lien avec le sens existant. Par ailleurs, une partie des sens créés font l'objet d'un entraînement pendant les 4 jours qui précèdent l'expérience, les autres non. Les auteurs de l'expérience observent que le temps de réponse est plus long quand le mot cible a fait l'objet d'un entraînement, ce qui indique que l'entraînement a conduit au stockage du sens nouveau dans le lexique mental des participants et que ce sens nouveau entre en compétition avec le sens préexistant pendant la tâche de décision. Parmi les mots pour lesquels un sens nouveau a été appris, le temps de réponse est plus long quand le nouveau sens n'est pas lié au sens déjà connu que quand il l'est, ce qui indique que les différents sens d'un mot entrent moins en compétition quand ils sont liés par polysémie que quand ils relèvent de l'homonymie.

2. L'évaluation de la régularité d'une polysémie

Le calcul du degré de régularité d'un patron de polysémie implique a minima l'évaluation de deux paramètres importants : i) la proportion des mots polysémiques relevant de ce patron par rapport à l'ensemble des mots susceptibles d'en relever et ii) la fréquence de ces mots polysémiques et la prise en compte de la distribution de leur sens (Lombard *et al.*, à paraître). Plus un patron a produit dans la langue de mots polysémiques par rapport à son potentiel génératif, plus il est régulier et il l'est d'autant plus que les mots qui en relèvent sont d'un usage fréquent et que les mots polysémiques qu'il a produits ont des sens équilibrés, autrement dit, de fréquence équivalente. On peut penser en effet que ces derniers sont de meilleurs représentants de la polysémie en question, en comparaison de ceux qui ont un sens source ou un sens cible rare. Si l'on prend l'exemple de la polysémie "action > objet résultant de cette action", les noms comme *construction* ou *traduction* devraient ainsi avoir plus de poids dans le calcul de la régularité de ce patron que des noms comme *assemblage*, dont le sens d'action est a priori beaucoup plus fréquent que le sens d'objet résultant. Dans la suite de cette section, nous commencerons par rendre compte d'études portant sur l'évaluation des effectifs des patrons de polysémie à partir des données disponibles dans les ressources lexicales (section 2.1) puis nous aborderons les quelques études s'intéressant à la modélisation de la polysémie régulière en corpus (section 2.2).

2.1. Évaluation de la régularité d'une polysémie à partir de lexiques

À la suite de l'étude de Buitelaar (1998) qui a le premier exploité la ressource lexicale de l'anglais WordNet (Miller *et al.*, 1990) pour en extraire des classes de mots présentant des alternances de sens similaires, plusieurs études ont proposé d'identifier des polysémies régulières en exploitant les chaînes d'hyponymes dans des ressources hiérarchiquement structurées (Buitelaar, 2000 ; Peters, 2006 ; Barque & Chaumartin, 2008 ; entre autres). Il résulte de ces travaux des inventaires de patrons de polysémie et des classes de mots instanciant ces patrons pour différentes langues. L'exploitation de WordNet donne accès pour l'anglais à un ensemble important de données, mais dont on ne peut évaluer la précision, celles-ci n'ayant pas fait l'objet d'une validation manuelle complète. L'étude de Buitelaar (2000) mentionne par exemple 1341 clusters de sens représentant potentiellement des extensions de sens régulières pour 3336 noms. Concernant le français, deux ressources

lexicales hiérarchisées⁵, complémentaires du point de vue de leur couverture et de leur précision, ont été exploitées avec le même objectif d'identifier les polysémies régulières et leurs effectifs (Barque *et al.*, 2018). Après validation manuelle des données extraites automatiquement, les auteurs de l'étude ont obtenu un ensemble de 50 patrons de type métonymique et 53 patrons de type métaphorique pour le français, dont respectivement 22 et 9 sont représentés dans les deux ressources.

Les données produites à partir de ressources lexicales ne peuvent toutefois suffire à une caractérisation précise du degré de régularité des patrons identifiés. En effet, les dictionnaires ne listent qu'une partie des mots de la langue et ne décrivent pas l'ensemble des sens possibles pour ces mots, notamment quand il s'agit de sens hautement prédictibles. Par exemple, seuls les noms qui dénotent des contenants servant prototypiquement de mesure voient leur sens cible décrit dans les dictionnaires. On y trouve ainsi plus fréquemment le sens cible de *verre* ('le contenu du verre') que celui de *flûte* ou de *mug*. Une autre limitation des ressources lexicales pour l'identification des patrons de polysémie régulière et de leurs effectifs concerne l'hétérogénéité des traitements en termes de distinction des sens et d'encodage des liens entre ces sens : il n'est pas rare que les alternances fortement régulières, comme la polysémie "action > résultat", soient décrites au sein d'une entrée tantôt par deux définitions distinctes (ex. le nom *construction* dans *Le Petit Robert* est défini d'une part comme A. 'action de construire' et B. 'ce qui est construit, bâti'), tantôt par une unique définition disjonctive (ex. le nom *affirmation* dans le même dictionnaire est défini comme 'action d'affirmer [...] ; jugement ainsi énoncé') sans qu'il apparaisse clairement si cette différence de traitement encode une différence linguistique entre polysémie et monosémie complexe (voir section 1).

2.2. Évaluation de la régularité d'une polysémie à partir de corpus

L'exploration de corpus constitue une autre façon de déterminer empiriquement l'inventaire des polysémies régulières d'une langue et de leurs instances. Cette méthode présente l'intérêt, par rapport à la méthode lexicographique, de donner des informations de fréquence sur les mots et leurs sens, second paramètre important pour le calcul du degré de régularité d'une polysémie. L'étude du phénomène en contexte permet par ailleurs de s'affranchir des questions de conventionnalisation des sens et d'accéder ainsi aussi bien aux sens systématiquement produits par un patron qu'aux productions discursives néologiques vouées ou non à intégrer

⁵ Ces deux ressources sont : le *Wolf* (Fiser & Sagot, 2015), construit automatiquement à partir du WordNet anglais, et le *Réseau lexical du français* ou *RL-fr* (Lux Pogodalla & Polguère, 2011), moins bruité puisque développé manuellement mais également moins couvrant que le *Wolf*.

le lexique (voir *L'omelette au jambon est partie sans payer*⁶). Il existe théoriquement deux cas de figure : soit on dispose de corpus sémantiquement annotés et donc des données nécessaires pour le calcul de la régularité des patrons au sein de ce corpus, soit on ne dispose que de corpus bruts, ce qui implique une étape préalable d'induction ou de repérage automatique d'alternances de sens récurrentes. La rareté des corpus annotés sémantiquement limite l'utilisation de la première méthode. Seul l'anglais dispose d'une telle ressource avec le corpus Sencor (Landes *et al.*, 1998), dans lequel environ 193 000 occurrences de noms, verbes et adjectifs ont été annotées manuellement d'après l'inventaire des sens de WordNet. La représentation de la polysémie régulière dans ce corpus n'a toutefois, à notre connaissance, pas encore fait l'objet d'une analyse approfondie. On peut citer toutefois des projets d'annotation plus circonscrits qui s'intéressent aux alternances sémantiques régulières (Martinez Alonso *et al.*, 2013 ; Jurgen, 2014), mais qui ont pour visée spécifique l'étude des cas d'ambiguïté contextuelle révélés par un désaccord entre annotateurs (ou par une annotation multiple, quand elle est permise dans le protocole d'annotation). L'ambiguïté contextuelle pouvant résulter de plusieurs phénomènes distincts (homonymie, polysémie – régulière ou non –, monosémie complexe, sous-détermination), les données produites dans le cadre de ces études ne peuvent être qu'en partie exploitées pour l'étude de la polysémie régulière.

Du côté de l'exploration de corpus non annotés sémantiquement, plusieurs études utilisent les représentations distribuées de mots pour évaluer leur appartenance ou non à une classe de polysémie régulière. Boleda *et al.* (2012) proposent un premier modèle de polysémie régulière qu'ils testent dans une tâche de classification binaire. Dans ce modèle, une alternance de sens A-B est représentée par la moyenne des représentations des classes sémantiques A et B, elles-mêmes respectivement représentées par la moyenne des représentations de mots monosémiques relevant de ces classes. La représentation des mots à classer, qui se limitent dans cette étude à des mots bisémiques dans la ressource WordNet, est calculée quant à elle à partir de l'ensemble de leurs occurrences. Le classement s'effectue par le calcul de la distance cosinus entre la représentation du mot à classer et celle de l'alternance de sens. La méthode, appliquée sur 60 alternances régulières issues de Corelex (Buitelaar, 1998), donne des résultats variés selon les alternances, ce que les auteurs expliquent par la proximité sémantique plus ou moins grande entre les deux sens en jeu dans l'alternance, proximité sémantique allant de pair avec une similarité distributionnelle variable. Del Tredici et Bel (2015) proposent un modèle qui s'inspire dans les grandes lignes du précédent mais qui s'en

⁶ Exemple emprunté à Nunberg (1978).

distingue en ce que les deux sens de l'alternance ne sont pas moyennés mais représentés indépendamment. Les mots polysémiques sont quant à eux représentés par leur degré d'appartenance à chacune des deux classes. Par exemple, les noms *bœuf* et *autruche*, dont les deux sens 'animal' et 'viande' sont lexicalisés, auront des indices distincts d'appartenance à chacune de ces deux classes, le sens 'viande' étant plus fréquent en corpus pour *bœuf* que pour *autruche*. Cette modélisation leur permet ainsi de rendre compte des différences de fréquence entre les sens du mot polysémique.

3. Les facteurs explicatifs de la polysémie régulière

Les polysémies régulières présentant des degrés de régularité variables, on peut se demander ce qui peut expliquer cette variation. On évoquera ici trois facteurs susceptibles d'influencer le degré de régularité d'un patron de polysémie : le type de figure sur laquelle repose l'extension de sens (section 3.1), la force du lien qui unit les deux sens en jeu (section 3.2) et enfin le caractère morphologiquement construit des mots qui instancient les sens impliqués dans un patron de polysémie (section 3.3).

3.1. Types d'extensions de sens

La figure lexicale sur laquelle repose une polysémie régulière peut s'imposer d'emblée comme l'un des facteurs explicatifs du degré de régularité de cette polysémie. Depuis (Apresjan, 1974), la régularité est en effet fréquemment décrite comme une propriété typique de la métonymie plutôt que de la métaphore⁷ et les études théoriques portant sur la polysémie régulière s'intéressent de fait majoritairement aux alternances fortement régulières de type métonymique (Nunberg & Zaenen, 1992 ; Pustejovsky, 1995 ; entre autres⁸). La corrélation entre figure et régularité est également postulée dans les études psycholinguistiques s'intéressant au traitement cognitif des différents types de mots à sens multiples et à leur représentation dans le lexique mental. Les travaux de Brocher *et al.* (2016, 2018), par exemple, montrent que les polysèmes irréguliers, représentés par des mots à sens cible métaphorique,

⁷ On peut interpréter cette idée de deux manières, non exclusive l'une de l'autre : soit il existe plus de polysémies régulières de types métonymiques que de polysémies régulières de types métaphoriques, soit les métonymies régulières ont tendance à avoir un degré de régularité supérieur à celui des métaphores régulières.

⁸ Voir aussi (Dölling, 2020) pour un inventaire des polysémies dites systématiques, toutes de type métonymique.

ont un traitement plus proche de celui des homonymes que les polysèmes réguliers, représentés par des mots à sens cible métonymique.

Le lien de causalité entre figure et régularité peut s'expliquer en partie par les propriétés structurelles des deux figures. La métonymie repose sur une relation de contiguïté entre le référent du sens dérivé et celui du sens source, ce qui implique la création d'un référent connecté à celui dénoté par le sens source. La métaphore, quant à elle, repose sur une relation d'analogie entre les deux référents, qui implique la création d'un référent disjoint du précédent. Cette différence implique qu'il y a tendanciellement une plus grande proximité entre deux sens liés par métonymie qu'entre deux sens liés par métaphore, notamment lorsque la contiguïté se traduit par une inclusion référentielle, c'est-à-dire lorsque le référent du sens dérivé correspond à une partie du référent du sens source (p. ex dans la métonymie "animal > viande"). Le sens cible est alors fortement prédictible. Ainsi, la production ou l'interprétation d'un sens métonymique, tendanciellement plus proche du sens source et plus prédictible qu'un sens produit par métaphore, demanderait aux locuteurs un effort cognitif moindre, ce qui expliquerait qu'il soit plus régulièrement produit.

Le traitement cognitif moins coûteux des métonymies par rapport à celui des métaphores a fait l'objet d'une validation empirique dans les travaux de Klepoustoniou *et al.* (2007, 2008, 2012). Les résultats des différentes expériences menées indiquent en effet (i) que les métonymies sont traitées significativement plus rapidement que les homonymes, et (ii) que les métaphores ont un traitement intermédiaire entre celui des métonymies et celui des homonymes. Pour ce qui est du lien entre figure et régularité à proprement parler, en revanche, il n'existe pas, à notre connaissance, d'étude d'envergure qui ait cherché à valider empiriquement la plus grande propension de la métonymie à la régularité, par rapport à la métaphore. En admettant que cela soit le cas, il reste que la figure ne peut expliquer à elle seule les degrés de régularité observés dans l'ensemble du lexique. En effet, comme l'ont souligné plusieurs auteurs (Copestake & Briscoe 1995 ; Pethö, 2007 ; Vincente & Falkum 2017 ; entre autres) et comme le montrent les inventaires de patrons de polysémie précédemment cités, les métaphores régulières ne sont pas rares et présentent, elles aussi, des degrés de régularité variables. Ces deux propriétés doivent donc absolument être décorrélées pour permettre une meilleure compréhension du phénomène (voir Lombard *et al.*, à paraître).

3.2. Degré de connectivité entre les deux sens

Le degré de connectivité entre deux sens – qui, selon la définition qu'on en donne, peut concerner différentes propriétés de la relation : similarité, associativité, prédictibilité – semble

être ainsi une variable explicative importante du degré de régularité de la polysémie en jeu. Plusieurs études typologiques s'intéressant à la polysémie font l'hypothèse d'un lien entre degré de connectivité des sens et régularité des cas de polysémie observés à travers les langues.

Dans l'étude de Srinivasan et Rabagliati (2015), des patrons de polysémie de l'anglais sont présentés à des locuteurs anglophones natifs d'une autre langue, qui doivent indiquer si les différents patrons proposés existent dans leur langue maternelle et, le cas échéant, si les mots polysémiques anglais relevant de ces patrons ont des équivalents dans leur langue. Les résultats de leur enquête montrent que la grande majorité des 27 alternances sémantiques régulières testées⁹ s'observent dans la plupart des langues considérées mais que ces patrons récurrents varient du point de vue de la taille de l'ensemble des instances comparables d'une langue à l'autre. Pour expliquer cette variation, les auteurs font l'hypothèse de contraintes plus ou moins fortes pesant sur le sens cible, qui détermineraient la prédictibilité de ce sens cible et donc sa fréquence de réalisation au sein des langues. Le sens de *viande*, par exemple, est plus prédictible à partir du sens d'animal (ex. *lapin*, *bœuf*) que le sens d'objet à partir du sens de matière (ex. *coton*, *verre*) et il est donc plus probable de trouver, d'une langue à l'autre, des instances similaires pour le premier patron que pour le second.

Dans une étude récente, Xu *et al.* (2020) s'intéressent de plus près aux causes des colexifications récurrentes à travers les langues¹⁰. Leur hypothèse est que la fréquence de colexification de deux concepts à travers les langues dépend de l'effort cognitif nécessaire à leur traitement, et que cet effort cognitif dépend lui-même du degré de connectivité entre les concepts en jeu : plus ces derniers sont connectés, plus leur traitement cognitif s'en trouve facilité et plus ils ont tendance à être colexifiés dans les langues. Le degré de connectivité est ici évalué à partir de mesures d'associativité et de mesures de similarité entre des mots représentant les concepts colexifiés¹¹. Ces deux modèles (associativité et similarité) s'avèrent

⁹ On notera que 4 des 27 patrons considérés dans cette étude relèvent du procédé morphologique de conversion et non de la polysémie à proprement parler, celle-ci n'impliquant pas, au contraire de la conversion, de changement de catégorie grammaticale. Ces patrons de conversion (ex. "instrument > se servir de cet instrument", *hammer*, etc.) s'avèrent être ceux qui sont le moins récurrents dans les langues étudiées. Du point de vue de la figure, les patrons testés sont dans leur très grande majorité métonymiques (24/27).

¹⁰ Sur la notion de colexification, qui recouvre en grande partie mais pas uniquement celle de polysémie (l'utilisation d'une même forme pour exprimer deux sens distincts s'appliquant également à l'homonymie), voir François (2008).

¹¹ Les mots représentant les concepts sont en anglais. Les mesures d'associativité sont calculées à partir des données d'un jeu en ligne dans lequel les participants doivent, à la lecture d'un mot cible, indiquer

être ceux qui prédisent le mieux les données de colexifications observées, extraites des *Intercontinental Dictionary Series* (Borin *et al.*, 2013), par rapport à deux autres modèles servant de baselines : l'un qui s'appuie sur des données de fréquence (moyenne des fréquences des deux mots représentant les concepts), l'autre qui s'appuie sur la notion de "métaphoricité" entre concepts, opérationnalisée par l'existence d'un changement de concrétude et/ou de polarité (perçu comme négatif vs positif) entre les deux concepts. Les résultats confirment ainsi l'hypothèse formulée par les auteurs selon laquelle les concepts qui ont la plus grande probabilité d'être associés et ceux qui apparaissent le plus souvent dans des contextes similaires sont ceux qui sont le plus souvent colexifiés dans les langues, autrement dit que le degré de connectivité entre deux concepts influence leur fréquence de colexification.

L'analyse des concepts associés aux mots polysémiques, dans l'étude qui vient d'être présentée, ne s'accompagne pas d'une généralisation en termes de patrons de polysémie. Par ailleurs, l'estimation qui est faite de la similarité entre concepts s'appuie sur des représentations distribuées de mots (*word embeddings*), dont les limites pour l'étude de la polysémie sont connues (Camacho-Collados & Pilehvar, 2018). L'utilisation de plus en plus fréquente de plongements contextualisés, qui représentent des occurrences de mots, permet d'envisager une analyse plus fine des rapports entre similarité sémantique et sens associés à une même forme lexicale. De fait, plusieurs études récentes cherchent à évaluer dans quelle mesure les modèles de langue contextuels captent la distinction entre polysémie et homonymie, et dans quelle mesure ces modèles coïncident avec les jugements humains sur le degré de similarité entre emplois d'un même mot (Nair *et al.*, 2020 ; Haber & Poesio, 2021 ; Trott & Bergen, 2021 ; entre autres). Bien que ne portant pas spécifiquement sur la question du degré de régularité d'une polysémie, les données présentées dans ces études nous donnent un premier aperçu des liens entre similarité distributionnelle et polysémie régulière. Elles confirment tout d'abord empiriquement que la polysémie occupe une position intermédiaire entre homonymie et monosémie. La mise en regard des résultats de ces études montre par ailleurs que la polysémie semble plus proche de la monosémie lorsque les cas étudiés relèvent tous de la polysémie régulière, comme dans les données de Haber et Poesio (2021)¹² que lorsque les mots polysémiques étudiés ne sont pas homogènes en termes de régularité, comme

le premier mot qui leur vient à l'esprit. Les mesures de similarité sont quant à elles calculées en corpus à partir de la distance entre les vecteurs des mots représentant les concepts.

¹² Il s'agit en l'occurrence de métonymies fortement régulières (par ex. "animal > viande", "contenant > contenu"). On notera que les auteurs incluent dans cet ensemble plusieurs cas de facettes (par ex. physical-information) qui influent certainement sur la proximité des jugements entre polysémie et monosémie.

dans les données de Trott et Bergen (2021). Les données de ces deux études ne sont toutefois bien sûr pas directement comparables et une expérimentation reste à mener sur la similarité distributionnelle entre sens de patrons de polysémie plus ou moins réguliers pour évaluer dans quelle mesure celle-ci explique ou non la régularité observée d'une polysémie.

3.3. Relations avec le lexique morphologiquement construit

Mentionnons enfin, comme dernier élément susceptible de nous donner des informations sur la régularité d'un patron de polysémie, les relations qu'entretiennent, au sein du lexique, les règles d'extension de sens avec les règles de construction morphologique. Deux aspects de ces relations sont à considérer : la rivalité entre procédés de création de sens, d'une part, et les propriétés morphologiques des mots polysémiques, d'autre part.

On peut faire l'hypothèse que, dans les langues, tous les sens construits à partir d'autres sens n'ont pas les mêmes probabilités d'être produits par une construction morphologique ou par une extension de sens. Certains types de sens vont avoir tendance à être exclusivement construits par la morphologie, comme la transposition, qui implique un changement de catégorie grammaticale que l'extension de sens seule ne peut opérer, ou encore l'antonymie. D'autres au contraire vont typiquement être construits par extension de sens comme c'est le cas des sens métaphoriques, qui reposent sur une analogie non exprimée avec le sens de base. Mais il existe également tout un ensemble de cas où construction morphologique et extension de sens entrent en concurrence (ex. Copestake & Briscoe, 1995). En français par exemple, il est possible de produire à partir d'un nom d'artefact (ex. *cuillère*, *table*) un nom ayant le sens de 'quantité' soit par dérivation sémantique (ex. *une cuillère de farine*, *une table de convives*) soit par dérivation morphologique à l'aide du suffixe *-ée* (ex. *une cuillerée de farine*, *une tablée de convives*). De même certains sens typiquement dérivés par polysémie dans une langue peuvent s'exprimer morphologiquement dans une autre langue, comme le montrent les données produites dans le cadre d'études typologiques (Zalizniak *et al.*, 2012 ; Zalizniak, 2018). La rivalité entre construction morphologique et extension de sens, qu'elle ait lieu au sein d'une langue ou entre langues distinctes, est ainsi susceptible de nous informer sur la régularité de ces dernières. Les constructions morphologiques reposant sur un principe de régularité formelle et sémantique, on peut faire l'hypothèse qu'une polysémie qui construit le même type de sens qu'une construction morphologique aura elle aussi tendance à être régulière. Par exemple, les noms d'agent, typiquement construits en *-eur* à partir d'un verbe (ex. *construire* > *constructeur*), peuvent aussi être produits par extension de sens à partir d'un

nom d'événement, que celui soit construit morphologiquement (ex. *accusation* 'action d'accuser' / 'partie qui accuse') ou non (ex. *censure* 'action de censurer' / 'groupe qui censure').

L'autre question est de savoir si le statut morphologique (construit ou non) des mots quiinstancient une polysémie est susceptible de nous informer sur la régularité de celle-ci. Les alternances de sens régulièrement observées au sein du lexique construit (Lehrer, 2003 ; Rainer, 2014) ont théoriquement trois origines possibles (Salvadori & Huyghe, ce numéro). La polysémie observée peut tout d'abord être héritée de la base morphologique, elle-même polysémique. Par exemple, les noms déverbaux dérivés de verbes présentant une alternance causative (ex. *sécher le linge* / *le linge sèche*) héritent souvent de cette alternance : *le séchage du linge* est ainsi ambigu entre une lecture causative et une lecture anticausative. Elle peut par ailleurs être le résultat d'une dérivation en parallèle à l'aide d'un suffixe polyvalent à partir d'une base sémantique unique. Le suffixe *-eur* par exemple permet de produire des noms d'animés et des noms d'instruments, ce qui donne lieu à des alternances récurrentes parmi les noms construits en *-eur* (ex. *distributeur*, qui peut dénoter une machine ou une personne qui distribue, avec la possibilité d'établir un lien d'analogie entre les deux). Elle peut enfin résulter d'une extension de sens prenant pour base le sens produit par la dérivation morphologique. Par exemple, la polysémie "action > groupe de personnes qui font l'action" de certains noms en *-ion* (ex. *accusation*, *opposition*) est le résultat d'une extension métonymique. Ces différents types de polysémies récurrentes observées au sein de classes lexicales morphologiquement construites font ainsi des propriétés morphologiques des mots polysémiques un élément à prendre en compte pour expliquer la régularité de ces derniers.

4. Présentation des contributions

Les quatre contributions regroupées dans ce numéro s'appuient sur l'exploration de ressources lexicales pour questionner certains des aspects de la polysémie régulière, tels que la possibilité de voir se réaliser un sens cible à partir d'un sens source, l'origine des polysémies observées au sein du lexique construit ou encore celle des polysémies récurrentes à travers les langues. Les trois premières contributions offrent une analyse de cas de polysémie en français liés à des classes circonscrites sémantiquement (noms de fleurs) et/ou morphologiquement (noms en *-isme* et noms déverbaux), tandis que la quatrième étude adopte une approche typologique en analysant un cas de polysémie complexe (c'est-à-dire correspondant à une série d'extensions de sens) récurrent dans les langues indo-européennes. Nous les présentons ici dans cet ordre.

L'étude présentée par **Alain Polguère** porte sur l'identification de la polysémie régulière dans le *Réseau lexical du français* (RL-fr), une ressource dans laquelle les unités lexicales, spécifiées sémantiquement, sont liées entre elles par différents types de relations (représentées par des fonctions lexicales), dont celle de co-polysémie. L'auteur y propose de décrire les polysémies régulières observées au sein d'une classe sémantique à partir de l'hyperonyme (accessible par la fonction lexicale *Gener*) des unités lexicales qui composent cette classe, ce qu'il illustre avec l'exemple des noms de fleurs en français. S'appuyant sur les descriptions du RL-fr, Alain Polguère montre ainsi que la définition lexicographique structurée de l'unité lexicale *fleur* encode les polysémies régulières observées pour les hyponymes de *fleur* (ex. *cyclamen*, *lavande*, *gentiane*, *lotus*). L'auteur illustre sa méthode de description lexicographique des patrons de polysémie régulière à l'aide de quatre patrons de la classe des noms de fleurs : (i) "fleur > couleur" (ex. *cyclamen*), (ii) "fleur > fragrance" (ex. *lavande*), (iii) "fleur > substance comestible" (ex. *gentiane*) et enfin (iv) "fleur > forme" (ex. *lotus*).

Grigory Agabalian s'intéresse aux alternances de sens régulièrement observées dans la classe morphologique des noms du français suffixés en *-isme*, dont certains peuvent dénoter une doctrine ou une attitude (ex. *hédonisme*, *scepticisme*). Après avoir proposé un ensemble de tests linguistiques révélateurs de l'une et l'autre de ces classes sémantiques et les avoir appliqués pour identifier les noms en *-isme* relevant conjointement des deux classes, Grigory Agabalian s'interroge sur la nature de la relation qui unit ces deux sens. Il montre que si les N de doctrine sont parfois dotés de facettes relevant d'une autre classe (ex des noms en *-isme* qui dénotent à la fois une doctrine et une activité, comme *chamanisme*), la relation entre les sens de doctrine et d'attitude sont clairement incompatibles et ne relèvent donc pas de la monosémie complexe. S'appuyant sur l'étude de descriptions issues du *Trésor de la Langue Française*, l'auteur montre ensuite que l'alternance a deux sources possibles : la polysémie, dans le cas où le sens attitude est dérivé du sens doctrine, comme c'est le cas pour *machiavélisme* par exemple ; et l'homonymie, dans le cas où les deux sens sont dérivés en parallèle lors de deux constructions morphologiques.

Justine Salvadori et **Richard Huyghe** s'intéressent à la question de l'origine des polysémies régulières s'observant dans le lexique construit, plus particulièrement dans l'ensemble des noms déverbaux. S'appuyant sur les descriptions d'un large ensemble de noms morphologiquement construits présentant des sens multiples (2954 noms déverbaux du français formés à l'aide de 46 suffixes), ils distinguent les cas où la polysémie régulièrement observée pour le nom construit est le fruit de deux dérivations morphologiques (avec lien de sens postulable a posteriori) des cas où elle résulte d'une seule dérivation morphologique avec

extension de sens. C'est ce second phénomène que les auteurs se proposent de décrire, en proposant la notion de patron morphologique complexe et en comparant les 40 patrons qu'ils observent au sein des noms déverbaux du français. Ils mettent ainsi en évidence des spécificités morphologiques associées à la polysémie régulière. Par exemple, parmi les suffixes produisant des noms d'événement (*-age*, *-ment*, *-ion*), le suffixe *-age* est le plus apte à produire des noms présentant une polysémie régulière de type 'événement' / 'artefact' (l'artefact étant le résultat de l'événement).

Enfin, dans une perspective typologique, **Mariia Orlova** propose une étude comparative d'une chaîne d'extension de sens s'appliquant à des verbes de mouvement dans 11 langues indo-européennes. L'auteure y décrit précisément les différentes étapes de dérivations permettant de passer du sens original de 'lever' (*to take up*) à celui de 'détruire' (*to destroy*) ou de 'abolir' (*to abolish*). S'appuyant sur les données de la base *SemShift* (Zalizniak, 2018) et sur celles de dictionnaires de référence des langues décrites, Mariia Orlova montre que cette polysémie s'observe dans les 11 langues étudiées, qu'elle est régulière dans chacune de ces langues, et enfin que la chaîne de glissements de sens comporte, selon les lexèmes et selon les langues, entre 2 et 5 sens attestés de la chaîne. L'étude met ainsi en lumière l'existence de patrons de polysémie régulière complexes récurrents à travers les langues et les structures cognitives qui leur sont sous-jacentes.

Bibliographie

- Apresjan, J. (1974). Regular polysemy. *Linguistics*, 14, 5-32.
- Asher, N. (2011). *Lexical Meaning in Context: A Web of Words*. Cambridge University Press.
- Barque L., & Chaumartin, F.-R. (2008). La polysémie régulière dans WordNet, In *Actes de la 15ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'2008)*, Avignon (France). (pp. 202-211). <http://talnarchives.atala.org/TALN/TALN-2008/index.html>
- Barque, L., Haas, P., & Huyghe, R. (2018). Polysémie régulière et néologie sémantique : constitution d'une ressource pour l'étude des sens nouveaux. *Neologica*, 12, 91-108.
- Boleda, G., Padó, S., & Utt, J. (2012). Regular polysemy: A distributional model. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics–Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)* (pp. 151-160). Association for Computational Linguistics.

- Borin, L., Comrie, B., & Saxena, A. (2013). The Intercontinental Dictionary Series – A rich and principled database for language comparison. In L. Borin & A. Saxena (Eds), *Approaches to measuring linguistic differences* (pp. 285-302). De Gruyter Mouton.
- Bréal, M. (1904). *Essai de sémantique (science des significations)*. Hachette.
- Brocher, A., Foraker, S., & Koenig, J. P. (2016). Processing of irregular polysemes in sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 1798–1813.
- Brocher, A., Koenig, J. P., Maurer, G., & Foraker, S. (2018). About sharing and commitment: the retrieval of biased and balanced irregular polysemes. *Language, Cognition and Neuroscience*, 33(4), 443-466.
- Buitelaar, P. (1998). *CoreLex: systematic polysemy and underspecification*. Phd Thesis, Brandeis University.
- Buitelaar, P. (2000). Reducing lexical semantic complexity with systematic polysemous classes and underspecification. In *NAACL-ANLP 2000 Workshop: Syntactic and Semantic Complexity in Natural Language Processing Systems* (Vol. 1, pp. 14-19). <https://doi.org.10.3115/1117543.1117546>
- Camacho-Collados, J., & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, 63, 743-788.
- Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of semantics*, 12(1), 15-67.
- Cruse, D. (1986). *Lexical Semantics*. Cambridge University Press.
- Cruse, D. (1995). Polysemy and related phenomena from a cognitive linguistic viewpoint. In P. St. Dizier & E. Viegas (Eds), *Computational Lexical Semantics* (pp.33-39). Cambridge University Press.
- Del Tredici, M., & Bel, N. (2015). A word-embedding-based sense index for regular polysemy representation. In P. Blunsom, S. Cohen, P. Dhillon & P. Liang (Eds), *Proceedings of the 1st workshop on vector space modeling for natural language processing* (pp. 70-78). Association for Computational Linguistics.
- Dölling, J. (2020). Systematic polysemy. In L. Matthewson, C. Meier, H. Rullmann, T. E. Zimmermann & D. Gutzmann (Eds), *The Wiley Blackwell Companion to Semantics*. Wiley Online Library. DOI: 10.1002/9781118788516.
- Falkum, I., & Vicente, A. (2015). Polysemy: Current perspectives and approaches. *Lingua*, 157, 1-16.

- Fiser, D., & Sagot, B. (2015). Constructing a poor man's wordnet in a resource-rich world. *Language Resources and Evaluation*, 49(3), 601-635.
- François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove (Ed.), *Studies in language companion series* (vol. 106, pp. 163–215). John Benjamins.
- Frisson, S. (2015). About bound and scary books: The processing of book polysemies. *Lingua*, 157, 17-35.
- Godard, D., & Jayez, J. (1996). Type nominaux et anaphores : le cas des objets et des événements. *Cahiers Chronos*, 1, 41-58.
- Haber, J., & Poesio, M. (2021). Patterns of Polysemy and Homonymy in Contextualised Language Models. In *Findings of the Association for Computational Linguistics: EMNLP 2021* (pp. 2663-2676). Association for Computational Linguistics.
- Jurgens, D. (2014). An analysis of ambiguity in word sense annotations. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland (pp. 3006-3012). European Language Resources Association (ELRA).
- Kleiber, G. (1999). *Problèmes de sémantique : la polysémie en questions*. Presses Universitaires du Septentrion.
- Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, 20(1), 1-24.
- Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and language*, 123(1), 11-21.
- Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1534-1543.
- Landes, S., Leacock, C., & Teng, R. (1998). Building semantic concordances. In C. Fellbaum (Ed.), *WordNet: An Electronic Lexical Database* (pp. 199-216). MIT Press.
- Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics*, 1/2, 207-246.
- Lehrer, A. (2003). Polysemy in derivational affixes. In B. Nerlich, Z. Todd, V. Herman & D. D. Clark (Eds), *Polysemy Flexible Patterns of Meaning in Mind and Language* (pp. 217-232). Walter de Gruyter.

- Lombard, A., Huyghe, R., Barque, L., & Gras, D. (In press) Regular polysemy and novel word-sense identification. *The Mental Lexicon*.
- Lux-Pogodalla, V., & Polguère, A. (2011). Construction of a French Lexical Network: Methodological Issues. In *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop, Ljubljana, Slovenia* (pp. 54-61). <https://hal.archives-ouvertes.fr/hal-00686467/document>
- Maciejewski, G., Rodd, J. M., Mon-Williams, M., & Klepousniotou, E. (2020). The cost of learning new meanings for familiar words. *Language, Cognition and Neuroscience*, 35(2), 188-210.
- Martinez Alonso, H., Pedersen, B. S., & Bel, N. (2013). Annotation of regular polysemy and underspecification. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (Vol. 2: Short Papers, pp. 725-730). Association for Computational Linguistics.
- Miller, G., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). WordNet: An online lexical database. *International Journal of Lexicography*, 3(4), 235-244.
- Moldovan A. (2021) Descriptions and tests for polysemy. *Axiomathes*, 31(3), 229-249.
- Nair S., Srinivasan M., & Meylan, S. (2020). Contextualized word embeddings encode aspects of human-like word sense knowledge. In *Proceedings of the Workshop on the Cognitive Aspects of the Lexicon* (pp 129–141). <https://doi.org/10.48550/arXiv.2010.13057>
- Nunberg, G. D. (1978). *The pragmatics of reference*. City University of New York.
- Nunberg, G., & Zaenen, A. (1992). Systematic Polysemy in Lexicology and Lexicography. In K.V. Hannu Tommola, T. Salmi-Tolonen & J. Schopp (Eds), *Proceedings of Euralex II* (pp. 387-396). University of Tampere.
- Ostler, N., & Atkins, S. (1991). Predictable Meaning Shift: Some Linguistic Properties of Lexical Implication Rules. In J. Pustejovsky & S. Bergler (Eds), *Lexical Semantics and Knowledge Representation: First SIGLEX Workshop Proceedings*. Springer-Verlag.
- Peters, W. (2003). Metonymy as a cross-lingual phenomenon. In *Proceedings of the ACL 2003 Workshop on Lexicon and Figurative Language* (pp. 1-9). Association for Computational Linguistics. <https://doi-org/10.3115/1118975.1118976>
- Peters, W. (2006). In Search for More Knowledge: Regular Polysemy and Knowledge Acquisition. In P. Sojka, K.-S. Choi, C. Fellbaum & P. Vossen (Eds.), *Proceedings of GWC 2006* (pp. 245–251). Masaryk University.

- Pethö, G. (2007). On Irregular Polysemy. In P. Marina, P. Gergely & R. Csilla (Eds.), *The cognitive basis of polysem: New sources of evidence for theories of word meaning*, Frankfurt (pp. 123-156). Peter Lang.
- Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press, Cambridge.
- Rainer, F. (2014). Polysemy in derivation. In R. Lieber & P. Štekauer (Eds), *The Oxford handbook of derivational morphology* (pp. 338–353). Oxford University Press.
- Ravin, Y., & Leacock, C. (2000). Polysemy: An Overview. In Y. Ravin & C. Leacock (Eds.), *Polysemy. Theoretical and Computational Approaches* (pp 1-29). Oxford University Press.
- Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, M. G., & Davis, M. H. (2012). Learning new meanings for old words: Effects of semantic relatedness. *Memory & Cognition*, 40(7), 1095-1108.
- Rodriguez, I. L. (2009). Of women, bitches, chickens and vixens: Animal metaphors for women in English and Spanish. *Cultura, lenguaje y representación: revista de estudios culturales de la Universitat Jaume I*, 77-100.
- Romeo, L., Mendes, S., & Bel, N. (2014). A cascade approach for complex type classification. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation – LREC'14* (pp. 4451–4458). European Language Resources Association (ELRA).
- Srinivasan, M., Berner, C., & Rabagliati, H. (2019). Children use polysemy to structure new word meanings. *Journal of Experimental Psychology: General*, 148(5), 926-942.
- Srinivasan, M., & Rabagliati, H. (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, 157, 124-152.
- Talebinejad, M. R., & Dastjerdi, H. V. (2005). A cross-cultural study of animal metaphors: When owls are not wise! *Metaphor and Symbol*, 20(2), 133-150.
- Trott, S., & Bergen, B. (2021). Raw-c: Relatedness of ambiguous words in context (a new lexical resource for English). In *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing*. <https://doi.org/10.48550/arXiv.2105.13266>
- Tuggy, D. (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics*, 4, 273–290.
- Victorri, B., & Fuchs, C. (1996). *La polysémie-Construction dynamique du sens*. Hermes.
- Xu, Y., Duong, K., Malt, B. C., Jiang, S., & Srinivasan, M. (2020). Conceptual relations predict colexification across languages. *Cognition*, 201, 104-280.
- Zalizniak, A., (2018). The catalogue of semantic shifts: 20 years later. *Russian journal of linguistics*, 22(4), 770-787.

- Zalizniak, A., Bulakh, M., Ganenkov, D., Gruntov, I., Maisak, T., & Russo, M. (2012). The catalogue of semantic shifts as a database for lexical semantic typology. *Linguistics*, 50(3), 633-669.
- Zwicky, A., & Sadock, J. (1975). Ambiguity tests and how to fail them. *Syntax and Semantics*, 4, 1-36.